

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS:

SCHMIDT, Harald
ZABEL, Ulrike
POLLER, Wolfgang

(ii) TITLE OF THE INVENTION: Isolated and purified human soluble
guanylyl cyclase $\alpha 1/\beta 1$ (hsGCa1/ $\beta 1$)

(iii) VOS-101

(iv) US 09/762,767
2001-02-13

(iv) PCT/DE99/02601
1999-08-16

(iv) DE 198 37 015.6
1998-08-14

(v) NUMBER OF SEQUENCES: 10

(vi) COMPUTER-READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) DATA TO SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3015 basepairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: doublestrand
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCCTTATGGC	GATTGGGCGG	CTGCAGAGAC	CAGGACTCAG	TTCCCCTGCC	CTAGTCTGAG	60
CCTAGTGGGT	GGGACTCAGC	TCAGAGTCAG	TTTTCCAGAA	GCAGGTTTCA	GTGCAGAGTT	120
TTCTCTACACT	TTTCCTGCGC	TAGAGCAGCG	AGCAGCCTGG	AACAGACCCA	GGCGGAGGAC	180
ACCTGTGGGG	GAGGGAGCGC	CTGGAGGAGC	TTAGAGACCC	CAGCCGGGCG	TGATCTCACC	240
ATGTGCGGAT	TTGCGAGGCG	CGCCCTGGAG	CTGCTAGAGA	TCCGGAAGCA	CAGCCCCGAG	300
GTGTGCGAAG	CCACCAAGAC	TGCGGCTCTT	GGAGAAAGCG	TGAGCAGGGG	GCCACCGCGG	360
TCTCCGCGCC	TGTCTGCACC	CTGTGCGCTG	AGCTGCCTGA	CAGTGACAAT	GACATCCCGAG	420
TTACCAAGTGT	CCTTGAATTG	ATAGTGGCTT	CTGTTTGTC	GTCTCATATA	AGAACTACAG	480
CTCATCAGGA	GGAGATCGCA	GCAGGGTAAG	AGACACCAAC	ACCATGTTCT	GCACGAAGCT	540
CAAGGATCTC	AAGATCACAG	GAGAGTGTCC	TTTCTCCTTA	CTGGCACCAG	GTCAAGTTCC	600
TAACGAGTCT	TCAGAGGAGG	CAGCAGGAAG	CTCAGAGAGC	TGCAAAGCAA	CCGTGCCCCT	660
CTGTCAAGAC	ATTCTTGAGA	AGAACATACA	AGAAAGTCTT	CCTCAAAGAA	AAACCAAGTCG	720
GAGCCGAGTC	TATCTTCACA	CTTTGGCAGA	GAGTATTTGC	AAACTGATTT	TCCCAGAGTT	780
TGAACGGCTG	AATGTTGCAC	TTTCAAGAAC	ATTGGCAAAG	CACAAAATAA	AAGAAAGCAG	840
GAAATCTTTG	GAAAGAGAAG	ACTTTGAAAA	AACAATTGCA	GAGCAAGCAG	TTGCAGCAGG	900
AGTTCCAGTG	GAGGTTATCA	AAGAATCTCT	TGGTGAAGAG	GTTTTTAAAA	TATGTTACGA	960
GGAAGATGAA	AACATCCTTG	GGGTGGTTGG	AGGCACCCTT	AAAGATTTTT	TAAACAGCTT	1020
CAGTACCCTT	CTGAAACAGA	GCAGCCATTG	CCAAGAAGCA	GGAAAAAGGG	GCAGGCTTGA	1080
GGACGCCTCC	ATTCTATGCC	TGGATAAGGA	GGATGATTTT	CTACATGTTT	ACTACTTCTT	1140
CCCTAAGAGA	ACCACCTCCC	TGATTCTTCC	CGGCATCATA	AAGGCAGCTG	CTCAGCTATT	1200
ATATGAAACG	GAAAGTGAAG	TGTCGTTAAT	GCCTCCCTGC	TTCCATAATG	ATTGCAGCGA	1260
GTTTGTGAAT	GAGCCCTACT	TGTTGTACTC	CGTTCACATG	AAAAGCACCA	AGCCATCCCT	1320
GTCCCCCAGC	AAACCCAGT	CCTCGCTGGT	GATTCCCACA	TCGCTATTCT	GCAAGACATT	1380

	TCCATTCCAT	TTCATGTTTG	ACAAAGATAT	GACAATTCTG	CAATTTGGCA	ATGGCATCAG	1440
	AAGGCTGATG	AACAGGAGAG	ACTTTCAAGG	AAAGCCTAAT	TTTGAAGAAT	ACTTTGAAAT	1500
	TCTGACTCCA	AAAATCAACC	AGACGTTTAG	CGGGATCATG	ACTATGTTGA	ATATGCAGTT	1560
	TGTTGTACGA	GTGAGGAGAT	GGGACAACTC	TGTGAAGAAA	TCTTCAAGGG	TTATGGACCT	1620
5	CAAAGGCCAA	ATGATCTACA	TTGTTGAATC	CAGTGCAATC	TTGTTTTTGG	GGTCACCCTG	1680
	TGTGGACAGA	TTAGAAGATT	TTACAGGACG	AGGGCTCTAC	CTCTCAGACA	TCCCAATTCA	1740
	CAATGCACTG	AGGGATGTGG	TCTTAATAGG	GGAACAAGCC	CGAGCTCAAG	ATGGCCTGAA	1800
	GAAGAGGCTG	GGGAAGCTGA	AGGCTACCCT	TGAGCAAGCC	CACCAAGCCC	TGGAGGAGGA	1860
	GAAGAAAAAG	ACAGTAGACC	TTCTGTGCTC	CATATTTCCC	TGTGAGGTTG	CTCAGCAGCT	1920
10	GTGGCAAGGG	CAAGTTGTGC	AAGCCAAGAA	GTTCAGTAAT	GTCACCATGC	TCTTCTCAGA	1980
	CATCGTTGGG	TTCACTGCCA	TCTGCTCCCA	GTGCTCACCG	CTGCAGGTCA	TCACCATGCT	2040
	CAATGCACTG	TACACTCGCT	TCGACCAGCA	GTGTGGAGAG	CTGGATGTCT	ACAAGGTGGA	2100
	GACCATTGGC	GATGCCTATT	GTGTAGCTGG	GGGATTACAC	AAAGAGAGTG	ATACTCATGC	2160
	TGTTTCAGTA	GCGCTGATGG	CCCTGAAGAT	GATGGAGCTC	TCTGATGAAG	TTATGTCTCC	2220
15	CCATGGAGAA	CCTATCAAGA	TGCGAATTGG	ACTGCACTCT	GGATCAGTTT	TTGCTGGCGT	2280
	CGTTGGAGTT	AAAATGCCCC	GTTACTGTCT	TTTTGGAAAC	AATGTCACCT	TGGCTAACAA	2340
	ATTTGAGTCC	TGCAGTGTA	CACGAAAAAT	CAATGTCAGC	CCAACAACCT	ACAGATTACT	2400
	CAAAGACTGT	CCTGGTTTCG	TGTTTACCCC	TCGATCAAGG	GAGGAACTTC	CACCAAACTT	2460
20	CCCTAGTGAA	ATCCCCGGAA	TCTGCCATTT	TCTGGATGCT	TACCAACAAG	GAACAACTC	2520
	AAAACCATGC	TTCCAAAAGA	AAGATGTGGA	AGATGGCAAT	GCCAATTTTT	TAGGCAAAGC	2580
	ATCAGGAATA	GATTAGCAAC	CTATATACCT	ATTTATAAGT	CTTTGGGGTT	TGACTCATTG	2640
	AAGATGTGTA	GAGCCTCTGA	AAGCACTTTA	GGGATTGTAG	ATGGCTAACA	AGCAGTATTA	2700
	AAATTTTCAGG	AGCCAAGTCA	CAATCTTTCT	CCTGTTTAA	ATGACAAAAT	GTACTCACTT	2760
	CAGTACTTCA	GCTCTTCAAG	AAAAAAAAAA	AAACCTTAAA	AAGCTACTTT	TGTGGGAGTA	2820
25	TTTCTATTAT	ATAACCAGCA	CTTACTACCT	GTACTCAAAA	TTCAGCACCT	TGTACATATA	2880
	TCAGATAATT	GTAGTCAATT	GTACAACTG	ATGGAGTCAC	CTGCAATCTC	ATATCCTGGT	2940
	GGAAATGCCAT	GGTTATTAAA	GTGTGTTTGT	GATAGTGTCG	TCAAAAAAAAA	AAAAAAAAAA	3000
	AAAAAAAAAA	AAAAA					3015

30

(2) DATA TO SEQ ID NO: 2:

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 695 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: protein (human soluble guanylyl cyclase a1 (hsGCal))

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

45

Met	Phe	Cys	Thr	Lys	Leu	Lys	Asp	Leu	Lys	Ile	Thr	Gly	Glu	Cys	Pro
				5					10					15	
Phe	Ser	Leu	Leu	Ala	Pro	Gly	Gln	Val	Pro	Asn	Glu	Ser	Ser	Glu	Glu
			20					25					30		
Ala	Ala	Gly	Ser	Ser	Glu	Ser	Cys	Lys	Ala	Thr	Val	Pro	Ile	Cys	Gln
		35					40				45				
50	Asp	Ile	Pro	Glu	Lys	Asn	Ile	Gln	Glu	Ser	Leu	Pro	Gln	Arg	Lys
		50				55					60				
Ser	Arg	Ser	Arg	Val	Tyr	Leu	His	Thr	Leu	Ala	Glu	Ser	Ile	Cys	Lys
		65			70				75					80	
Leu	Ile	Phe	Pro	Glu	Phe	Glu	Arg	Leu	Asn	Val	Ala	Leu	Gln	Arg	Thr
			85					90					95		
Leu	Ala	Lys	His	Lys	Ile	Lys	Glu	Ser	Arg	Lys	Ser	Leu	Glu	Arg	Glu
		100					105					110			
Asp	Phe	Glu	Lys	Thr	Ile	Ala	Glu	Gln	Ala	Val	Ala	Ala	Gly	Val	Pro
		115				120					125				
60	Val	Glu	Val	Ile	Lys	Glu	Ser	Leu	Gly	Glu	Glu	Val	Phe	Lys	Ile
		130				135						140			

	Tyr	Glu	Glu	Asp	Glu	Asn	Ile	Leu	Gly	Val	Val	Gly	Gly	Thr	Leu	Lys	
	145					150					155					160	
	Asp	Phe	Leu	Asn	Ser	Phe	Ser	Thr	Leu	Leu	Lys	Gln	Ser	Ser	His	Cys	
				165						170					175		
5	Gln	Glu	Ala	Gly	Lys	Arg	Gly	Arg	Leu	Glu	Asp	Ala	Ser	Ile	Leu	Cys	
				180					185					190			
	Leu	Asp	Lys	Glu	Asp	Asp	Phe	Leu	His	Val	Tyr	Tyr	Phe	Phe	Pro	Lys	
			195					200					205				
10	Arg	Thr	Thr	Ser	Leu	Ile	Leu	Pro	Gly	Ile	Ile	Lys	Ala	Ala	Ala	His	
	210						215					220					
	Val	Leu	Tyr	Glu	Thr	Glu	Val	Glu	Val	Ser	Leu	Met	Pro	Pro	Cys	Phe	
	225					230					235					240	
	His	Asn	Asp	Cys	Ser	Glu	Phe	Val	Asn	Gln	Pro	Tyr	Leu	Leu	Tyr	Ser	
				245					250						255		
15	Val	His	Met	Lys	Ser	Thr	Lys	Pro	Ser	Leu	Ser	Pro	Ser	Lys	Pro	Gln	
				260					265					270			
	Ser	Ser	Leu	Val	Ile	Pro	Thr	Ser	Leu	Phe	Cys	Lys	Thr	Phe	Pro	Phe	
				275				280					285				
20	His	Phe	Met	Phe	Asp	Lys	Asp	Met	Thr	Ile	Leu	Gln	Phe	Gly	Asn	Gly	
	290						300					305					
	Ile	Arg	Arg	Leu	Met	Asn	Arg	Arg	Asp	Phe	Gln	Gly	Lys	Pro	Asn	Phe	
	310					315					320					325	
	Glu	Glu	Tyr	Phe	Glu	Ile	Leu	Thr	Pro	Lys	Ile	Asn	Gln	Thr	Phe	Ser	
				330						335					340		
25	Gly	Ile	Met	Thr	Met	Leu	Asn	Met	Gln	Phe	Val	Val	Arg	Val	Arg	Arg	
				345				350						355			
	Trp	Asp	Asn	Ser	Val	Lys	Lys	Ser	Ser	Arg	Val	Met	Asp	Leu	Lys	Gly	
			360					365					370				
30	Gln	Met	Ile	Tyr	Ile	Val	Glu	Ser	Ser	Ala	Ile	Leu	Phe	Leu	Gly	Ser	
	375						380					385					
	Pro	Cys	Val	Asp	Arg	Leu	Glu	Asp	Phe	Thr	Gly	Arg	Gly	Leu	Tyr	Leu	
	390					395					400					405	
	Ser	Asp	Ile	Pro	Ile	His	Asn	Ala	Leu	Arg	Asp	Val	Val	Leu	Ile	Gly	
					410					415					420		
35	Glu	Gln	Ala	Arg	Ala	Gln	Asp	Gly	Leu	Lys	Lys	Arg	Leu	Gly	Lys	Leu	
				425				430					435				
	Lys	Ala	Thr	Leu	Glu	Gln	Ala	His	Gln	Ala	Leu	Glu	Glu	Glu	Lys	Lys	
			440					445				450					
40	Lys	Thr	Val	Asp	Leu	Leu	Cys	Ser	Ile	Phe	Pro	Cys	Glu	Val	Ala	Gln	
	455						460					465					
	Gln	Leu	Trp	Gln	Gly	Gln	Val	Val	Gln	Ala	Lys	Lys	Phe	Ser	Asn	Val	
	470					475					480					485	
	Thr	Met	Leu	Phe	Ser	Asp	Ile	Val	Gly	Phe	Thr	Ala	Ile	Cys	Ser	Gln	
				490					495						500		
45	Cys	Ser	Pro	Leu	Gln	Val	Ile	Thr	Met	Leu	Asn	Ala	Leu	Tyr	Thr	Arg	
				505				510						515			
	Phe	Asp	Gln	Gln	Cys	Gly	Glu	Leu	Asp	Val	Tyr	Lys	Val	Glu	Thr	Ile	
			520					525					530				
50	Gly	Asp	Ala	Tyr	Cys	Val	Ala	Gly	Gly	Leu	His	Lys	Glu	Ser	Asp	Thr	
	535					540					545						
	His	Ala	Val	Gln	Ile	Ala	Leu	Met	Ala	Leu	Lys	Met	Met	Glu	Leu	Ser	
	550					555					560					565	
	Asp	Glu	Val	Met	Ser	Pro	His	Gly	Glu	Pro	Ile	Lys	Met	Arg	Ile	Gly	
				570					575						580		
55	Leu	His	Ser	Gly	Ser	Val	Phe	Ala	Gly	Val	Val	Gly	Val	Lys	Met	Pro	
				585				590						595			
	Arg	Tyr	Cys	Leu	Phe	Gly	Asn	Asn	Val	Thr	Leu	Ala	Asn	Lys	Phe	Glu	
			600					605					610				
60	Ser	Cys	Ser	Val	Pro	Arg	Lys	Ile	Asn	Val	Ser	Pro	Thr	Thr	Tyr	Arg	
	615					620					625						
	Leu	Leu	Lys	Asp	Cys	Pro	Gly	Phe	Val	Phe	Thr	Pro	Arg	Ser	Arg	Glu	

10

15

20

25

30

35

40

45

50

55

60

TTACCTAACA	TGGTGATCTG	CAAGTAGTAG	GCACCCAATA	AATATTTGTT	GAATTTAGTT	2280
AAATGAAACT	GAACAGTGTT	TGGCCATGTG	TATATTTATA	TCATGTTTAC	CAAATCTGTT	2340
TAGTGTTCCA	CATATATGTA	TATGTATATT	TTAATGACTA	TAATGTAATA	AAGTTTATAT	2400
CATGTTGGTG	TATATCATTA	TAGAAATCAT	TTTCTAAAGG	AGT		2443

5

(2) DATA TO SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 619 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein (human soluble guanylyl cyclase b1 (hsGCb1))

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

20	Met	Tyr	Gly	Phe	Val	Asn	His	Ala	Leu	Glu	Leu	Leu	Val	Ile	Arg	Asn
					5					10					15	
	Tyr	Gly	Pro	Glu	Val	Trp	Glu	Asp	Ile	Lys	Lys	Glu	Ala	Gln	Leu	Asp
				20					25					30		
25	Glu	Glu	Gly	Gln	Phe	Leu	Val	Arg	Ile	Ile	Tyr	Asp	Asp	Ser	Lys	Thr
			35					40					45			
	Tyr	Asp	Leu	Val	Ala	Ala	Ala	Ser	Lys	Val	Leu	Asn	Leu	Asn	Ala	Gly
		50					55					60				
	Glu	Ile	Leu	Gln	Met	Phe	Gly	Lys	Met	Phe	Phe	Val	Phe	Cys	Gln	Glu
		65			70						75				80	
30	Ser	Gly	Tyr	Asp	Thr	Ile	Leu	Arg	Val	Leu	Gly	Ser	Asn	Val	Arg	Glu
					85					90				95		
	Phe	Leu	Gln	Asn	Leu	Asp	Ala	Leu	His	Asp	His	Leu	Ala	Thr	Ile	Tyr
			100						105					110		
35	Pro	Gly	Met	Arg	Ala	Pro	Ser	Phe	Arg	Cys	Thr	Asp	Ala	Glu	Lys	Gly
			115					120					125			
	Lys	Gly	Leu	Ile	Leu	His	Tyr	Tyr	Ser	Glu	Arg	Glu	Gly	Leu	Gln	Asp
		130					135					140				
	Ile	Val	Ile	Gly	Ile	Ile	Lys	Thr	Val	Ala	Gln	Gln	Ile	His	Gly	Thr
		145				150					155				160	
40	Glu	Ile	Asp	Met	Lys	Val	Ile	Gln	Gln	Arg	Asn	Glu	Glu	Cys	Asp	His
				165						170					175	
	Thr	Gln	Phe	Leu	Ile	Glu	Glu	Lys	Glu	Ser	Lys	Glu	Glu	Asp	Phe	Tyr
			180					185						190		
45	Glu	Asp	Leu	Asp	Arg	Phe	Glu	Glu	Asn	Gly	Thr	Gln	Glu	Ser	Arg	Ile
		195						200					205			
	Ser	Pro	Tyr	Thr	Phe	Cys	Lys	Ala	Phe	Pro	Phe	His	Ile	Ile	Phe	Asp
		210					215					220				
	Arg	Asp	Leu	Val	Val	Thr	Gln	Cys	Gly	Asn	Ala	Ile	Tyr	Arg	Val	Leu
		225				230					235				240	
50	Pro	Gln	Leu	Gln	Pro	Gly	Asn	Cys	Ser	Leu	Leu	Ser	Val	Phe	Ser	Leu
				245						250					255	
	Val	Arg	Pro	His	Ile	Asp	Ile	Ser	Phe	His	Gly	Ile	Leu	Ser	His	Ile
			260					265						270		
55	Asn	Thr	Val	Phe	Val	Leu	Arg	Ser	Lys	Glu	Gly	Leu	Leu	Asp	Val	Glu
		275						280					285			
	Lys	Leu	Glu	Cys	Glu	Asp	Glu	Leu	Thr	Gly	Thr	Glu	Ile	Ser	Cys	Leu
		290				295						300				
	Arg	Leu	Lys	Gly	Gln	Met	Ile	Tyr	Leu	Pro	Glu	Ala	Asp	Ser	Ile	Leu
		305			310						315				320	
60	Phe	Leu	Cys	Ser	Pro	Ser	Val	Met	Asn	Leu	Asp	Asp	Leu	Thr	Arg	Arg
				325						330					335	

Gly Leu Tyr Leu Ser Asp Ile Pro Leu His Asp Ala Thr Arg Asp Leu
 340 345 350
 Val Leu Leu Gly Glu Gln Phe Arg Glu Glu Tyr Lys Leu Thr Gln Glu
 355 360 365
 5 Leu Glu Ile Leu Thr Asp Arg Leu Gln Leu Thr Leu Arg Ala Leu Glu
 370 375 380
 Asp Glu Lys Lys Lys Thr Asp Thr Leu Leu Tyr Ser Val Leu Pro Pro
 385 390 395 400
 10 Ser Val Ala Asn Glu Leu Arg His Lys Arg Pro Val Pro Ala Lys Arg
 405 410 415
 Tyr Asp Asn Val Thr Ile Leu Phe Ser Gly Ile Val Gly Phe Asn Ala
 420 425 430
 Phe Cys Ser Lys His Ala Ser Gly Glu Gly Ala Met Lys Ile Val Asn
 435 440 445
 15 Leu Leu Asn Asp Leu Tyr Thr Arg Phe Asp Thr Leu Thr Asp Ser Arg
 450 455 460
 Lys Asn Pro Phe Val Tyr Lys Val Glu Thr Val Gly Asp Lys Tyr Met
 465 470 475 480
 20 Thr Val Ser Gly Leu Pro Glu Pro Cys Ile His His Ala Arg Ser Ile
 485 490 495
 Cys His Leu Ala Leu Asp Met Met Glu Ile Ala Gly Gln Val Gln Val
 500 505 510
 Asp Gly Glu Ser Val Gln Ile Thr Ile Gly Ile His Thr Gly Glu Val
 515 520 525
 25 Val Thr Gly Val Ile Gly Gln Arg Met Pro Arg Tyr Cys Leu Phe Gly
 530 535 540
 Asn Thr Val Asn Leu Thr Ser Arg Thr Glu Thr Thr Gly Glu Lys Gly
 545 550 555 560
 30 Lys Ile Asn Val Ser Glu Tyr Thr Tyr Arg Cys Leu Met Ser Pro Glu
 565 570 575
 Asn Ser Asp Pro Gln Phe His Leu Glu His Arg Gly Pro Val Ser Met
 580 585 590
 Lys Gly Lys Lys Glu Pro Met Gln Val Trp Phe Leu Ser Arg Lys Asn
 595 600 605
 35 Thr Gly Thr Glu Glu Thr Lys Gln Asp Asp Asp
 610 615

(2) DATA TO SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 14 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide (amino acids 634-647 of hsGCal)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Phe Thr Pro Arg Ser Arg Glu Glu Leu Pro Pro Asn Phe Pro
 5 10

(2) DATA TO SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide (amino acids 593-614 of hsGcb1)

(xi) SEQUENCEDESCRIPTION: SEQ ID NO: 6:

5 Lys Gly Lys Lys Glu Pro Met Gln Val Trp Phe Leu Ser Arg Lys Asn
 5 10 15
Thr Gly Thr Glu Glu Thr
 20

10 (2) DATA TO SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 28 basepairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single strans
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

20 (iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

25 (xi) SEQUENCEDESCRIPTION: SEQ ID NO: 7:

AAAAGGATCC ATGTTCTGCA CGAAGCTC

28

30

2) DATA TO SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 18 basepairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single strans
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

40

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

45

(xi) SEQUENCEDESCRIPTION: SEQ ID NO: 8:

ATTATGGAAG CAGGGAGG

50

18

2) DATA TO SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 28 basepairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single strans
(D) TOPOLOGY: linear

60

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

5 (xi) SEQUENCEDESCRIPTION: SEQ ID NO: 9:

AAAAGGATCC ATGTACGGAT TTGTGAAT

28

10 2) DATA TO SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 basepairs

(B) TYPE: nucleotide

15 (C) STRANDEDNESS: single strans

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

20 (iii) HYPOTHETICAL: no

(iv) ANTISENSE: no

25 (xi) SEQUENCEDESCRIPTION: SEQ ID NO: 10:

ATGCCGTGATT CCTGGGTACC

20